

WHAT IS CLAIMED IS:

Sub  
ai

1. A method of determining a background intensity an image comprising:  
selecting a plurality of spots within the image falling within a least squares curve fit;  
and  
determining a constant background intensity for the spots within the curve fit.

2. The method of Claim 1, further comprising determining a ratio an experimental image to a control image.

3. The method of Claim 2, further comprising determining the least squares curve fit from the equation:

$$r_m = R(g_m - g_b) + r_b = Rg_m + k$$

where  $r_m$  and  $g_m$  are the measured values of the images,  $r_b$ , and  $g_b$  are the background intensities of the images, and  $k$  is a constant.

4. The method of Claim 3, further comprising applying a constraint so the background intensities are greater than the bias levels.

5. The method of Claim 3, further comprising applying a constraint so the background intensities create a zero intercept of a linear regression of the equation:

$$(r_m - r_b) = m(g_m - g_b) + b$$

Equation 13

such that  $b$  is approximately zero, which occurs when

$$b = mg_b - r_b.$$

6. The method of Claim 5, further comprising extracting the background subtraction constants .

7. A method of selecting a microarray scan for analysis comprising:

determining a coefficient of variation for the microarray scan;

3 comparing the coefficient of variation to a predetermined threshold; and  
 4 selecting a microarray scan if the coefficient of variation is lower than the  
 5 predetermined threshold.

1 8. The method of Claim 7, further comprising determining the coefficient of  
 2 variation from the equation:

$$CV = \frac{R_{SD}}{R}$$

$$\text{where } R_{SD} \approx \sqrt{g_{SD}^2 \frac{\bar{r}^2}{\bar{g}^4} + \frac{r_{SD}^2}{\bar{g}^2} - 2\sigma_{rg} \frac{\bar{r}}{\bar{g}^3}}.$$

1 9. The method of Claim 7, further comprising determining a spot coefficient of  
 2 variation.

1 10. The method of Claim 7, further comprising determining an average coefficient  
 2 of variation.

1 11. A method of extracting data from an image comprising:

2 determining a covariance and a variance the of the image;

3 normalizing the covariance;

4 determining the average and standard deviation of the covariance; and

5 selecting the data based on the average and standard deviation of the covariance.

1 12. The method of Claim 11, further comprising calculating the covariance  
 2 according to the following equation:

$$\sigma_{rg} = \frac{1}{n} \sum_{i=1}^n (r_i - \bar{r})(g_i - \bar{g}).$$

1 13. The method of Claim 11, further comprising normalizing the covariance by  
 2 adding the variances in quadrature according to the following equation:

$$\sigma'_{rg} = \frac{\sigma_{rg}}{\sqrt{\sigma_r^2 + \sigma_g^2}},$$

where  $\sigma'_{rg}$  is the normalized covariance, and  $\sigma_r$ , and  $\sigma_g$  are the variances of the control and experimental channels.

14. The method of Claim 11, further comprising normalizing the covariance by adding the variances according to the following equation:

$$\sigma'_{rg} = \frac{\sigma_{rg}}{\left[ \frac{(\sigma_r + \sigma_g)}{2} \right]},$$

where  $\sigma'_{rg}$  is the normalized covariance, and  $\sigma_r$ , and  $\sigma_g$  are the variances of the control and experimental channels.

15. The method of Claim 11, further comprising normalizing the covariance by using a control channel variance according to the following equation:

$$\sigma'_{rg} = \frac{\sigma_{rg}}{\sigma_g},$$

where  $\sigma'_{rg}$  is the normalized covariance, and  $\sigma_r$ , and  $\sigma_g$  are the variances of the control and experimental channels.

16. The method of Claim 11, further comprising normalizing the covariance by using an experimental channel variance according to the following equation

$$\sigma'_{rg} = \frac{\sigma_{rg}}{\sigma_r},$$

where  $\sigma'_{rg}$  is the normalized covariance, and  $\sigma_r$ , and  $\sigma_g$  are the variances of the control and experimental channels.

17. A method of extracting data from an image comprising:

2 determining a covariance and a variance the of the image;

3 determining the slope of the covariance plotted against the variance; and

4 selecting the data where the slope exceeds a predetermined threshold.

1 18. The method of Claim 17, further comprising plotting each covariance value  
2 versus the average variance values.

1 19. The method of Claim 17, further comprising ignoring data points not along  
2 the slope of the covariance plotted against the variance.

1 20. The method of Claim 17, further comprising performing linear regression of  
2 the covariance plotted against the variance to create a distribution of data points.

1 21. The method of Claim 20, further comprising selecting an image having a tight  
2 distribution of data points.

**ABSTRACT**

Expression profiling using DNA microarrays is an important new method for analyzing cellular physiology. In "spotted" microarrays, fluorescently labeled cDNA from experimental and control cells is hybridized to arrayed target DNA and the arrays imaged at two or more wavelengths. Statistical analysis is performed on microarray images and show that non-additive background, high intensity fluctuations across spots, and fabrication artifacts interfere with the accurate determination of intensity information. The probability density distributions generated by pixel-by-pixel analysis of images can be used to measure the precision with which spot intensities are determined. Simple weighting schemes based on these probability distributions are effective in improving significantly the quality of microarray data as it accumulates in a multi-experiment database. Error estimates from image-based metrics should be one component in an explicitly probabilistic scheme for the analysis of DNA microarray data.

10075388.doc